

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

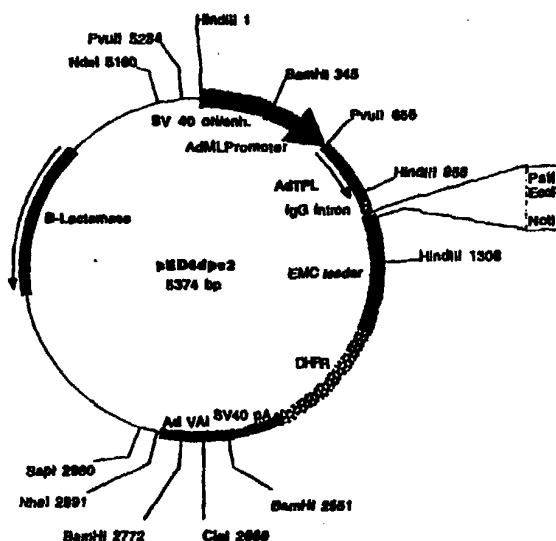
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/12, C07K 14/47, A61K 38/17</b>		<b>A2</b>	(11) International Publication Number: <b>WO 98/32853</b>
			(43) International Publication Date: 30 July 1998 (30.07.98)
(21) International Application Number: PCT/US98/01396		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 23 January 1998 (23.01.98)		<p><b>Published</b></p> <p><i>Without international search report and to be republished upon receipt of that report.</i></p>	
(30) Priority Data: 08/788,789      24 January 1997 (24.01.97)      US			
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).			
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).			
(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).			

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

## (57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED04pc2  
Plasmid size: 5374 bp

Comments/References: pED04pc2 is derived from pED04pc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al. (1991), NAR 19: 4485-4490.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5

10

## SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This application is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/788,789), filed January 24, 1997, which is incorporated by reference herein.

20

### FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

25

### BACKGROUND OF THE INVENTION

30

35

40

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA35\_2 deposited under accession number ATCC 98303;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA35\_2 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643; the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765; the nucleotide sequence of the full-length protein coding sequence of clone AA35\_2 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone AA35\_2 deposited

under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein  
5 comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group  
10 consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
- (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32;
- (c) fragments of the amino acid sequence of SEQ ID NO:2; and
- 15 (d) the amino acid sequence encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.

20 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID  
25 NO:3 from nucleotide 71 to nucleotide 736;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343;
- 30 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303;

5 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

10 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

15 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 71 to nucleotide 736; the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343; the nucleotide sequence of the full-length protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303. In yet other preferred  
20  
25 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising  
30 a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;

- (c) fragments of the amino acid sequence of SEQ ID NO:4; and
- (d) the amino acid sequence encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such  
5 protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence  
of SEQ ID NO:4 from amino acid 1 to amino acid 91.

In one embodiment, the present invention provides a composition comprising an  
isolated polynucleotide selected from the group consisting of:

- 10 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID  
NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID  
NO:5 from nucleotide 55 to nucleotide 423;
- (c) a polynucleotide comprising the nucleotide sequence of the full-  
length protein coding sequence of clone BG137\_7 deposited under accession  
15 number ATCC 98303;
- (d) a polynucleotide encoding the full-length protein encoded by the  
cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;
- (e) a polynucleotide comprising the nucleotide sequence of the mature  
protein coding sequence of clone BG137\_7 deposited under accession number  
20 ATCC 98303;
- (f) a polynucleotide encoding the mature protein encoded by the  
cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding a protein comprising the amino acid  
sequence of SEQ ID NO:6;
- 25 (h) a polynucleotide encoding a protein comprising a fragment of the  
amino acid sequence of SEQ ID NO:6 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of  
(a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein  
30 of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions  
to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID  
NO:5 from nucleotide 55 to nucleotide 423; the nucleotide sequence of the full-length

protein coding sequence of clone BG137\_7 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone BG137\_7 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of  
5 clone BG137\_7 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ  
10 ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:6;
- 15 (b) the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123;
- (c) fragments of the amino acid sequence of SEQ ID NO:6; and
- (d) the amino acid sequence encoded by the cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;

20 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 25 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID  
30 NO:7 from nucleotide 873 to nucleotide 2030;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173;



(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030; the nucleotide sequence of SEQ ID NO:7 from nucleotide 873 to nucleotide 2030; the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173; the nucleotide sequence of the full-length protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- 5 (b) the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329;
- (c) fragments of the amino acid sequence of SEQ ID NO:8; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;
- 10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 156 to nucleotide 677;
- 20 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;
- 25 (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;
- 30 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

5 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677; the nucleotide sequence of SEQ ID NO:10 from nucleotide 156 to nucleotide 677; the nucleotide sequence of the full-length protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303; or the  
10 nucleotide sequence of the mature protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303. In yet other preferred  
15 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:10, SEQ ID NO:9 or SEQ ID NO:12 .

20 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 120 to  
25 amino acid 189;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such  
30 protein comprises the amino acid sequence of SEQ ID NO:11 or the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755;

5 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631;

10 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP111\_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;

15 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP111\_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;

20 (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

25 (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755; the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631; the nucleotide sequence of the full-length protein coding sequence of clone CP111\_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CP111\_1 deposited

under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2011 to nucleotide 2565;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760; the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760; the nucleotide sequence of SEQ ID NO:15 from nucleotide 2011 to nucleotide 2565; the nucleotide sequence of the full-length protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
- (b) the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784;

(c) fragments of the amino acid sequence of SEQ ID NO:16; and

- 5 (d) the amino acid sequence encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784.

- 10 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

- 15 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147;

- 20 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;

- 25 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;

- 30 (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

5 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074; the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074; the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147; the nucleotide sequence of the full-length protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) fragments of the amino acid sequence of SEQ ID NO:18; and
- (c) the amino acid sequence encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;

25 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820;



(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN1120\_2 deposited under accession number ATCC 98303;

5 (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN1120\_2 deposited under accession number ATCC 98303;

10 (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;

15 (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

20 (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820; the nucleotide sequence of the full-length protein coding sequence of clone DN1120\_2 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone  
25 DN1120\_2 deposited under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid  
30 1 to amino acid 61.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- 5 (b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61;
- (c) fragments of the amino acid sequence of SEQ ID NO:20; and
- (d) the amino acid sequence encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;
- 10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 649 to nucleotide 786;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 736 to nucleotide 786;
- 20 (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO589\_1 deposited under accession number ATCC 98303;
- 25 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO589\_1 deposited under accession number ATCC 98303;
- 30 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:22 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

5 (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID  
10 NO:21 from nucleotide 649 to nucleotide 786; the nucleotide sequence of SEQ ID NO:21 from nucleotide 736 to nucleotide 786; the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787; the nucleotide sequence of the full-length protein coding sequence of clone DO589\_1 deposited under accession number ATCC 98303; or the nucleotide sequence of the mature protein coding sequence of clone DO589\_1 deposited  
15 under accession number ATCC 98303. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:21.

20 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:22;

(b) fragments of the amino acid sequence of SEQ ID NO:22; and

25 (c) the amino acid sequence encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:22.

In certain preferred embodiments, the polynucleotide is operably linked to an  
30 expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

- 5 The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which  
10 specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

15

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

20

#### DETAILED DESCRIPTION

##### ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone  
25 in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have  
30 determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation

proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

5        Clone "AA35\_2"

A polynucleotide of the present invention has been identified as clone "AA35\_2". AA35\_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer  
10    analysis of the amino acid sequence of the encoded protein. AA35\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AA35\_2 protein").

The nucleotide sequence of AA35\_2 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the  
15    predicted amino acid sequence of the AA35\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AA35\_2 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for AA35\_2 was searched against the  
20    GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AA35\_2 demonstrated at least some similarity with sequences identified as C16789 (Human placenta cDNA 5'-end GEN-529D11), H23653 (yn72e01.r1 Homo sapiens cDNA clone 173976 5' similar to contains Alu repetitive element), L31848 (Homo sapiens serine/threonine kinase receptor 2 (SKR2) gene, 3 alternative splices, 3'  
25    ends), U40455 (Human chromosome X cosmid, clones 196B12, 9H11 and 43H9, repeat units and sequence tagged sites), and Z82197 (Human DNA sequence from clone J293L6). The predicted amino acid sequence disclosed herein for AA35\_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AA35\_2 protein demonstrated at least some similarity to sequences  
30    identified as U58658 (unknown [Homo sapiens]) and X55777 (put. ORF [Homo sapiens]). Based upon sequence similarity, AA35\_2 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of AA35\_2 indicates that it may contain an Alu repetitive element.

Clone "AM42\_3"

A polynucleotide of the present invention has been identified as clone "AM42\_3". AM42\_3 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
5 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AM42\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AM42\_3 protein").

The nucleotide sequence of AM42\_3 as presently determined is reported in SEQ  
10 ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the AM42\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 2 to 14 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 15, or are a transmembrane domain.

15 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AM42\_3 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for AM42\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AM42\_3 demonstrated at least some similarity with sequences  
20 identified as AA109637 (mm01f02.r1 Stratagene mouse kidney (#937315) Mus musculus cDNA clone 520251 5'), AA131170 (zo08e05.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567104 3'), AA131483 (zo08e05.r1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567104 5'), and AA445683 (vf62h07.r1 Barstead MPLRB1 Mus musculus cDNA clone 848413 5'). Based upon  
25 sequence similarity, AM42\_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the AM42\_3 protein sequence centered around amino acid 152 of SEQ ID NO:4.

Clone "BG137\_7"

30 A polynucleotide of the present invention has been identified as clone "BG137\_7". BG137\_7 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. BG137\_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG137\_7 protein").

The nucleotide sequence of BG137\_7 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BG137\_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG137\_7 should be approximately 500 bp.

The nucleotide sequence disclosed herein for BG137\_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG137\_7 demonstrated at least some similarity with sequences identified as D87683 (Human mRNA for KIAA0243 gene, partial cds). Based upon sequence similarity, BG137\_7 proteins and each similar protein or peptide may share at least some activity.

#### Clone "CH699\_1"

A polynucleotide of the present invention has been identified as clone "CH699\_1". CH699\_1 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CH699\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CH699\_1 protein").

The nucleotide sequence of CH699\_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CH699\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 217 to 229 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 230, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CH699\_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for CH699\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. CH699\_1 demonstrated at least some similarity with sequences identified as AA155014 (mr99h05.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone 605625 5'), AA423476 (ve76d07.r1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone 832141 5'), U79271 (Human clones 23920 and 5 23921 mRNA sequence), and W72147 (zd70f08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 346023 3'). The predicted amino acid sequence disclosed herein for CH699\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CH699\_1 protein demonstrated at least some similarity to sequences identified as X51591 (beta-myosin heavy chain [Homo 10 sapiens]). Based upon sequence similarity, CH699\_1 proteins and each similar protein or peptide may share at least some activity.

#### Clone "CO851\_1"

A polynucleotide of the present invention has been identified as clone "CO851\_1". 15 CO851\_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO851\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein 20 as "CO851\_1 protein").

The nucleotide sequence of the 5' portion of CO851\_1 as presently determined is reported in SEQ ID NO:9. An additional internal nucleotide sequence from CO851\_1 as presently determined is reported in SEQ ID NO:10. What applicants believe is the proper reading frame and the predicted amino acid sequence encoded by such internal sequence 25 is reported in SEQ ID NO:11. Amino acids 3 to 15 of SEQ ID NO:11 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 16, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of CO851\_1, including the polyA tail, is reported in SEQ ID NO:12.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone 30 CO851\_1 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for CO851\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO851\_1 demonstrated at least some similarity with sequences identified as AA132585 (zo20c04.r1 Stratagene colon (#937204) Homo sapiens cDNA clone



587430 5'), H51262 (yp83b07.s1 Homo sapiens cDNA clone 194005 3'), W44070 (mc73a09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 354136 5'), and X92871 (X.laevis mRNA for an unknown transmembrane protein). The predicted amino acid sequence disclosed herein for CO851\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO851\_1 protein demonstrated at least some similarity to sequences identified as X92871 (unknown transmembrane protein [Xenopus laevis]). Based upon sequence similarity, CO851\_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CO851\_1 indicates that it may contain an Alu repetitive element.

#### Clone "CP111\_1"

A polynucleotide of the present invention has been identified as clone "CP111\_1". CP111\_1 was isolated from a human adult salivary gland cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CP111\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CP111\_1 protein").

The nucleotide sequence of CP111\_1 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CP111\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 40 to 52 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 53, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CP111\_1 should be approximately 3200 bp.

The nucleotide sequence disclosed herein for CP111\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CP111\_1 demonstrated at least some similarity with sequences identified as T53688 (ya98g07.r1 Homo sapiens cDNA clone 69756 5') and W70295 (zd58f03.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344861 3'). The predicted amino acid sequence disclosed herein for CP111\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol.

The predicted CP111\_1 protein demonstrated at least some similarity to sequences identified as X88852 (env protein [Primate T-cell lymphotropic]). Based upon sequence similarity, CP111\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the CP111\_1 protein sequence centered around amino acid 50 of SEQ ID NO:14.

#### Clone "CS278\_1"

A polynucleotide of the present invention has been identified as clone "CS278\_1". CS278\_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CS278\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CS278\_1 protein").

The nucleotide sequence of CS278\_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CS278\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 52 to 64 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 65, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CS278\_1 should be approximately 4400 bp.

The nucleotide sequence disclosed herein for CS278\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CS278\_1 demonstrated at least some similarity with sequences identified as AA234319 (zr66c07.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 668364 5'), H44192 (yo73f09.r1 Homo sapiens cDNA clone 183593 5'), W18258 (mb86a11.r1 Soares mouse p3NMF19), X76589 (H.sapiens DNA 3' flanking simple sequence region clone wg2c3), and Z74652 (M.musculus mRNA; expressed sequence tag (tcc2)). The predicted amino acid sequence disclosed herein for CS278\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CS278\_1 protein demonstrated at least some similarity to sequences identified as M34651 (ORF-3 protein [Suid herpesvirus 1]). The predicted CS278\_1 protein also demonstrated at least some similarity to a protein motif, cytochrome P450 cysteine heme-

iron ligand signature. Based upon sequence similarity, CS278\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts five potential transmembrane domains within the CS278\_1 protein sequence, which are centered around amino acids 75, 160, 525, 610, and 700 of SEQ ID NO:16, respectively. The nucleotide sequence of CS278\_1 may contain GAA simple repeat elements.

#### Clone "DF968\_3"

A polynucleotide of the present invention has been identified as clone "DF968\_3". DF968\_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DF968\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DF968\_3 protein").

The nucleotide sequence of DF968\_3 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DF968\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain. Another possible DF968\_3 reading frame and predicted amino acid sequence is encoded by basepairs 191 to 430 of SEQ ID NO:17 and is reported in SEQ ID NO:33.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DF968\_3 should be approximately 1010 bp.

The nucleotide sequence disclosed herein for DF968\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DF968\_3 demonstrated at least some similarity with sequences identified as AA426010 (zw49e12.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 773422 3' similar to contains element LTR5 repetitive element), H18256 (yn48a04.r1 Homo sapiens cDNA clone 171630 5'), and T06820 (EST04709 Homo sapiens cDNA clone HFBDZ29). The predicted amino acid sequence disclosed herein for DF968\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted DF968\_3 protein demonstrated at least some

similarity to sequences identified as Z38125 (orf, len 112, CAI 0.07). Based upon sequence similarity, DF968\_3 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of DF968\_3 indicates that it may contain repeat sequences.

5

#### Clone "DN1120\_2"

A polynucleotide of the present invention has been identified as clone "DN1120\_2". DN1120\_2 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
10 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DN1120\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DN1120\_2 protein").

The nucleotide sequence of DN1120\_2 as presently determined is reported in SEQ  
15 ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DN1120\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DN1120\_2 should be approximately 1000 bp.

20 The nucleotide sequence disclosed herein for DN1120\_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DN1120\_2 demonstrated at least some similarity with sequences identified as M62256 (EST00323 Homo sapiens cDNA clone HHCH15 similar to Alu repetitive element), M78991 (EST01139 Homo sapiens cDNA clone HHCPG39), Q59179  
25 (Human brain Expressed Sequence Tag EST00323), and Q61084 (Human brain Expressed Sequence Tag EST01139). Based upon sequence similarity, DN1120\_2 proteins and each similar protein or peptide may share at least some activity.

#### Clone "DO589\_1"

30 A polynucleotide of the present invention has been identified as clone "DO589\_1". DO589\_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DO589\_1 is a full-length

clone, including the entire coding sequence of a secreted protein (also referred to herein as "DO589\_1 protein").

The nucleotide sequence of DO589\_1 as presently determined is reported in SEQ ID NO:21. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DO589\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:22. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DO589\_1 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for DO589\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DO589\_1 demonstrated at least some similarity with sequences identified as AA402420 (zu47e04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 741150 3'), AA426621 (zw03a09.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 768184 5'), AA436749 (zv67c10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 758706 5'), H12845 (yj14h06.r1 Homo sapiens cDNA clone 148763 5'), R42350 (yg01b05.s1 Homo sapiens cDNA clone 30909 3'), W02775 (zc65g07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 327228 3'), W24833 (zc65g07.r1 Soares fetal heart), W58173 (zd19f02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 341115 3' similar to contains Alu repetitive element; contains element L1 repetitive element), and Z82201 (Human DNA sequence from clone J345P10). Based upon sequence similarity, DO589\_1 proteins and each similar protein or peptide may share at least some activity.

#### Deposit of Clones

Clones AA35\_2, AM42\_3, BG137\_7, CH699\_1, CO851\_1, CP111\_1, CS278\_1, DF968\_3, DN1120\_2, and DO589\_1 were deposited on January 23, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98303, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited

by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*,  
 5 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be  
 10 isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite  
 15 deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below,  
 20 and should be most reliable in isolating the clone of interest.

<u>Clone</u>	<u>Probe Sequence</u>
AA35_2	SEQ ID NO:23
AM42_3	SEQ ID NO:24
25 BG137_7	SEQ ID NO:25
CH699_1	SEQ ID NO:26
CO851_1	SEQ ID NO:27
CP111_1	SEQ ID NO:28
CS278_1	SEQ ID NO:29
30 DF968_3	SEQ ID NO:30
DN1120_2	SEQ ID NO:31
DO589_1	SEQ ID NO:32

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- (b) It should be designed to have a  $T_m$  of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with  $\gamma$ - $^{32}\text{P}$  ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately  $4\text{e}+6$  dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100  $\mu\text{l}$  of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100  $\mu\text{g}/\text{ml}$ . The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100  $\mu\text{g}/\text{ml}$  and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100  $\mu\text{g}/\text{ml}$  of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to  $1\text{e}+6$  dpm/mL. The filter is then preferably

incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The  
5 filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis,  
10 hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, *Bio/Technology* 10, 773-778 (1992) and in R.S.  
15 McDowell, *et al.*, *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion  
20 could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence  
25 listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

30 The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited



to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed  
5 sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s)  
10 corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-  
15 39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce  
20 gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through  
25 deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* 14(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination,  
30 preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for

the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>‡</sup>	Hybridization Temperature and Buffer <sup>†</sup>	Wash Temperature and Buffer <sup>†</sup>
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T <sub>B</sub> <sup>*</sup> ; 1xSSC	T <sub>B</sub> <sup>*</sup> ; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T <sub>D</sub> <sup>*</sup> ; 1xSSC	T <sub>D</sub> <sup>*</sup> ; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T <sub>F</sub> <sup>*</sup> ; 1xSSC	T <sub>F</sub> <sup>*</sup> ; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T <sub>H</sub> <sup>*</sup> ; 4xSSC	T <sub>H</sub> <sup>*</sup> ; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T <sub>J</sub> <sup>*</sup> ; 4xSSC	T <sub>J</sub> <sup>*</sup> ; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T <sub>L</sub> <sup>*</sup> ; 2xSSC	T <sub>L</sub> <sup>*</sup> ; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T <sub>N</sub> <sup>*</sup> ; 6xSSC	T <sub>N</sub> <sup>*</sup> ; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T <sub>P</sub> <sup>*</sup> ; 6xSSC	T <sub>P</sub> <sup>*</sup> ; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T <sub>R</sub> <sup>*</sup> ; 4xSSC	T <sub>R</sub> <sup>*</sup> ; 4xSSC

<sup>‡</sup> The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed

to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

- 5 1: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

- 10 \*T<sub>B</sub> - T<sub>R</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(°C) = 81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = 0.165 M).

- 15 Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

- 20 Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing  
25 polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

- The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein  
30 recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed  
35 by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205

cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as

those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and  
5 subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or  
10 all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic  
15 animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are  
20 known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic  
25 compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications  
30 of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art

(see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

### **USES AND BIOLOGICAL ACTIVITY**

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

#### **Research Uses and Utilities**

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that

described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

5 The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

20 Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

#### 25 Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.



Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors  
5 discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3,  
10 MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those  
15 described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, *Immunologic studies in Humans*); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology*  
20 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a.  
25 Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon  $\gamma$ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine  
30 Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature* 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in*

- Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991;
- 5 Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

- Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring
- 10 proliferation and cytokine production) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al.,
- 15 *Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger et al., *Eur. J. Immun.* 11:405-411, 1981; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988.

#### Immune Stimulating or Suppressing Activity

- 20 A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well
- 25 as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses,
- 30 herpesviruses, mycobacteria, *Leishmania* spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the

molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to  
5 anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

10 The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as  
15 described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

20 Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms.  
25 Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from  
30 the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and

murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy.

- 5 Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B  
10 lymphocyte antigens systemically.

- Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune  
15 responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a  
20 costimulatory signal to, and thereby activate, T cells *in vivo*.

- In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present  
25 invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The  
30 transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary

costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a  
5 cytoplasmic-domain truncated portion) of an MHC class I  $\alpha$  chain protein and  $\beta_2$  microglobulin protein or an MHC class II  $\alpha$  chain protein and an MHC class II  $\beta$  chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated  
10 immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated  
15 immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without  
20 limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al.,  
25 J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowman et al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al.,  
30 Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro*

antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnolli et al., *J. Immunol.* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., *J. Immunol.* 134:536-544, 1995; Inaba et al., *Journal of Experimental Medicine* 173:549-559, 1991; Macatonia et al., *Journal of Immunology* 154:5071-5079, 1995; Porgador et al., *Journal of Experimental Medicine* 182:255-260, 1995; Nair et al., *Journal of Virology* 67:4062-4069, 1993; Huang et al., *Science* 264:961-965, 1994; Macatonia et al., *Journal of Experimental Medicine* 169:1255-1264, 1989; Bhardwaj et al., *Journal of Clinical Investigation* 94:797-807, 1994; and Inaba et al., *Journal of Experimental Medicine* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., *Cytometry* 13:795-808, 1992; Gorczyca et al., *Leukemia* 7:659-670, 1993; Gorczyca et al., *Cancer Research* 53:1945-1951, 1993; Itoh et al., *Cell* 66:233-243, 1991; Zacharchuk, *Journal of Immunology* 145:4037-4045, 1990; Zamai et al., *Cytometry* 14:891-897, 1993; Gorczyca et al., *International Journal of Oncology* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., *Blood* 84:111-117, 1994; Fine et al., *Cellular Immunology* 155:111-122, 1994; Galy et al., *Blood* 85:2770-2778, 1995; Toki et al., *Proc. Nat. Acad. Sci. USA* 88:7548-7551, 1991.

### Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell

lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. *Cellular Biology* 15:141-151, 1995; Keller et al., *Molecular and Cellular Biology* 13:473-486, 1993; McClanahan et al., *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., *Proc. Natl. Acad. Sci. USA* 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and



Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long  
5 term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

10

#### Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns,  
15 incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as  
20 well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal  
25 disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue  
30 destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in

circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

#### Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin  $\alpha$  family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-

$\beta$  group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

#### Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion

- include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. 5 APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

#### Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity.

- 10 As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting 15 therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

- Assay for hemostatic and thrombolytic activity include, without limitation, those 20 described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### Receptor/Ligand Activity

- 25 A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, 30 cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without

limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- 5        Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987;
- 10    Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

#### Anti-Inflammatory Activity

- 15        Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or
- 20    suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin
- 25    lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

#### Cadherin/Tumor Invasion Suppressor Activity

30        Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human

diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved  
5 extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only  
10 itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with  
15 polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion  
20 suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

25 Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and  
30 reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from

forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

- 5           Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present
- 10          invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

- Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995;
- 15          Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

#### Tumor Inhibition Activity

- In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities.
- 20          A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating
- 25          or inhibiting factors, agents or cell types which promote tumor growth.

#### Other Activities

- A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious
- 30          agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms;



effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

#### 15 ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects

of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active

ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a  
5 therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines,  
10 lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with  
15 cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or  
20 cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical  
25 composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils  
30 may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein

of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such

antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in

5 R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where

10 abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage,

15 tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or

20 tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the

25 composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

30 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and

polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaVallie, Edward R.  
Racie, Lisa A.  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM

(iii) NUMBER OF SEQUENCES: 33

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genetics Institute, Inc.  
(B) STREET: 87 CambridgePark Drive  
(C) CITY: Cambridge  
(D) STATE: MA  
(E) COUNTRY: U.S.A.  
(F) ZIP: 02140

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Sprunger, Suzanne A.  
(B) REGISTRATION NUMBER: 41,323

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 498-8284  
(B) TELEFAX: (617) 876-5851

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1433 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear



(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCCCGTGGTT ACACAGCTAA TAGGTGGTGG AGATGGAGAC AGAATTCAAA CCCAGGCATT	60
CTTGATCTAC AGTATACACT CTTACCCACC ATCCTACACA GCCTTTCTTA TTCATAAAAT	120
ATTTTCTACA GTGCAAGAAA ATTTTGATAG CTTGCTTATT TATTCAAGAT TTAGACTATA	180
TAGATTAACT AGACTATCAA GATTTTAAAT TCTTGTGTTT TTTGTTTTTY YCCCCCTCTG	240
TGGCATAACT ATCTCTTAGT GATTTGAAGT TCTGATAGGC ATTTATTTAT GTTTTTGATT	300
AATTAAAAAA AGGGAAAAAA ATGGAACATA ATTATTGAAG CTATCGTCTA GGTAAAAACC	360
TTTCTAAATG TAAGGTTTCAT TTAGATTGAT GACCTGTAGA GTGTAACAGT ATTGCCATAG	420
GCATACAGCT TTTTAATCAC ATATCATACA TAAACAAATT AGTAATACAG GTGGGTAGAT	480
ACAGACCCTA ACTTTGAGCT CTAAGATGAA ATTTGTTTTAT AAATCCCTAG TTTCCATTCA	540
GT'TTTTCAA TATTTATCAA ACACCTACTG TGCCAGGCAT TGT'TTAGGCA CAGGGGATAC	600
AGCAGGAGAA CAAAATGAAC AAAATTTTTT GCCTTCACAG AGCTAATTTT TTGTATTTTT	660
TTGTAGAGAT GGGGTTTTGC CATGTTTGCC AGTCTGGTCT CAACCTCCTA AGCTCAAGCA	720
GCCCACCCTC CTTGGCTTCC CAAAGTGCTG AGATTACAGG CATGAGCCAC CGCACTCTTC	780
TTAGCTATTT TTCATAGAAA CTTTATGTAT AAAAATAGAA GGGTAATGAC ACACCACCTT	840
TCTACTGATC TCCCCACTTC AGTAGTTATC ACATAACAGT CTTTTTTCAC CTATCTCCTT	900
CACTTTACCT CCTCTCCCTT AGTACTTTGA AGTAAATCTC AATGCAAGCT GGTATGTTTT	960
TCAAAATGAA ACATATAAAC ATGGACTAGA AAAAAATCTC TTCATACAGG ATTTGGTTTT	1020
GCAGAGAATT TACAAAGTGC GGTTAATGTA TGCCAATGGT TTCTCAGTTT GGATATCGAG	1080
ATCCTTAGAT GGACCATGAA GCTGGTAATA ATTTTATAGC TAACTTTTGT TAAGTGCTTA	1140
CTATATGCCA GGCAGTGTTC TAAGCATTTT ACGTGTATTC ATTCATTCAG TTCTCACAAC	1200
TCTTTTAATT AGGTATTATT ATGATCTCCA TCTCAAAACA AAACAAAACA AAAAAATTAG	1260
CCTGGCATGG TGGCAGGCGC CTGTAATCCC AGTTACTTGA GAGGCTAAGG CAGGAGAATC	1320
GCTTGAATCT GGGAGGCAGA GGTTCAGTG AGCCGAGATT GCACTACTGC ACTCCAGCCT	1380
GGGTGACAGA ATGAGACTCT GTCTCAAAAA AAAAAAAAAA AAAAAAAAAA AAA	1433

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Phe Val Tyr Lys Ser Leu Val Ser Ile Gln Phe Phe Gln Tyr
 1             5             10             15

Leu Ser Asn Thr Tyr Cys Ala Arg His Cys Leu Gly Thr Gly Asp Thr
      20             25             30

Ala Gly Glu Gln Asn Glu Gln Asn Phe Leu Pro Ser Gln Ser
      35             40             45

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1401 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

TCGGGACAGA TTAAAGTGCA GCGTGGATTT TTTTTTCTC ACTTGCCTT GTGTTTTCCA      60
CCCTGAAAGA ATGTTGTGGC TGCTCTTTTT TCTGGTGA CTGAACTCTG      120
TCAACCAGGT GCAGAAAATG CTTTAAAGT GAGACTTAGT ATCAGAACAG CTCTGGGAGA      180
TAAAGCATAT GCCTGGGATA CCAATGAAGA ATACCTCTTC AAAGCGATGG TAGCTTTCTC      240
CATGAGAAAA GTTCCCAACA GAGAAGCAAC AGAAATTTC CATGTCCTAC TTTGCAATGT      300
AACCCAGAGG GTATCATTCT GGTTTGTGGT TACAGACCCT TCAAAAAATC ACACCCTTCC      360
TGCTGTTGAG GTGCAATCAG CCATAAGAAT GAACAAGAAC CGGATCAACA ATGCCTTCTT      420
TGTAATGAC CAACTCTGG AATTTTAAA AATCCCTTC AACTTGCAC CACCCATGGA      480

```

```

CCCATCTGTG CCCATCTGGA TTATTATATT TGGTGTGATA TTTTGCATCA TCATAGTTGC      540
AATTGCACTA CTGATTTTAT CAGGGATCTG GCAACGTAGA AGAAAGAACA AAGAACCATC      600
TGAAGTGGAT GACGCTGAAG ATAAGTGTGA AAACATGATC ACAATTGAAA ATGGCATCCC      660
CTCTGATCCC CTGGACATGA AGGGAGGGCA TATTAATGAT GCCTTCATGA CAGAGGATGA      720
GAGGCTCACC CCTCTCTGAA GGGCTGTTGT TCTGCTTCCT CAAGAAATTA AACATTTGTT      780
TCTGTGTGAC TGCTGAGCAT CCTGAAATAC CAAGAGCAGA TCATATATTT TGTTCACCA      840
TTCTTCTTTT GTAATAAATT TTGAATGTGC TTGAAAGTGA AAAGCAATCA ATTATACCCA      900
CCAACACCAC TGAAATCATA AGCTATTAC GACTCAAAAT ATTCTAAAT ATTTTCTGA      960
CAGTATAGTG TATAAATGTG GTCATGTGGT ATTTGTAGTT ATTGATTTAA GCATTTTGTAG    1020
AAATAAGATC AGGCATATGT ATATATTTTC ACACCTCAAA GACCTAAGGA AAAATAAATT    1080
TTCCAGTGGG GAATACATAT AATATGGTGT AGAAATCATT GAAATGGAT CCTTTTGTGAC    1140
GATCACTTAT ATCACTCTGT ATATGACTAA GTAAACAAAA GTGAGAAGTA ATTATTGTAA    1200
ATGGATGGAT AAAAATGGAA TTACTCATAT ACAGGGTGGG ATTTTATCCT GTTATCACAC    1260
CAACAGTTGA TTATATATTT TCTGAATATC AGCCCTAAT AGGACAATTC TATTGTGTA    1320
CCATTTCTAC AATTTGTAAA AGTCCAATCT GTGCTAACTT AATAAAGTAA TAATCATCTC    1380
TTTTAAAAAA AAAAAAAAAA A                                             1401

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala Glu Leu
1             5             10             15

Cys Gln Pro Gly Ala Glu Asn Ala Phe Lys Val Arg Leu Ser Ile Arg
                20             25             30

Thr Ala Leu Gly Asp Lys Ala Tyr Ala Trp Asp Thr Asn Glu Glu Tyr
35             40             45

```

Leu Phe Lys Ala Met Val Ala Phe Ser Met Arg Lys Val Pro Asn Arg  
 50 55 60  
 Glu Ala Thr Glu Ile Ser His Val Leu Leu Cys Asn Val Thr Gln Arg  
 65 70 75 80  
 Val Ser Phe Trp Phe Val Val Thr Asp Pro Ser Lys Asn His Thr Leu  
 85 90 95  
 Pro Ala Val Glu Val Gln Ser Ala Ile Arg Met Asn Lys Asn Arg Ile  
 100 105 110  
 Asn Asn Ala Phe Phe Val Asn Asp Gln Thr Leu Glu Phe Leu Lys Ile  
 115 120 125  
 Pro Ser Thr Leu Ala Pro Pro Met Asp Pro Ser Val Pro Ile Trp Ile  
 130 135 140  
 Ile Ile Phe Gly Val Ile Phe Cys Ile Ile Ile Val Ala Ile Ala Leu  
 145 150 155 160  
 Leu Ile Leu Ser Gly Ile Trp Gln Arg Arg Arg Lys Asn Lys Glu Pro  
 165 170 175  
 Ser Glu Val Asp Asp Ala Glu Asp Lys Cys Glu Asn Met Ile Thr Ile  
 180 185 190  
 Glu Asn Gly Ile Pro Ser Asp Pro Leu Asp Met Lys Gly Gly His Ile  
 195 200 205  
 Asn Asp Ala Phe Met Thr Glu Asp Glu Arg Leu Thr Pro Leu  
 210 215 220

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 441 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGGCCG CAG GTCTAGAATT CAATCGGCCA CAAGCTACTC TTTGGAGCCC ATCTATGGTT 60  
 TGTGGTATGA CCACTCCTCC AACTTCTCCT GGAAATGTCC CACCTGATCT GTCACACCCT 120  
 TACAGTAAAG TCTTTGGTAC AACTGCAGGT GGAAAAGGAA CTCCTCTGGG AACCCAGCA 180  
 ACCTCTCCTC CTCCAGCCCC ACTCTGTCAT TCGGATGACT ACGTGCACAT TTCACTCCCC 240

CAGGCCACAG TCACACCCCC CAGGAAGGAA GAGAGAATGG ATTCTGCAAG ACCATGTCTA 300  
CACAGACAAC ACCATCTTCT GAATGACAGA GGATCAGAAG AGCCACCTGG CAGCAAAGGT 360  
TCTGTCACTC TAAGTGATCT TCCAGGGTTT TTAGGTGATC TGGCCTCTGA AGAAGATAGT 420  
ATTGAAAAAA AAAAAAAAAA A 441

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Val Cys Gly Met Thr Thr Pro Pro Thr Ser Pro Gly Asn Val Pro  
1 5 10 15  
Pro Asp Leu Ser His Pro Tyr Ser Lys Val Phe Gly Thr Thr Ala Gly  
20 25 30  
Gly Lys Gly Thr Pro Leu Gly Thr Pro Ala Thr Ser Pro Pro Pro Ala  
35 40 45  
Pro Leu Cys His Ser Asp Asp Tyr Val His Ile Ser Leu Pro Gln Ala  
50 55 60  
Thr Val Thr Pro Pro Arg Lys Glu Glu Arg Met Asp Ser Ala Arg Pro  
65 70 75 80  
Cys Leu His Arg Gln His His Leu Leu Asn Asp Arg Gly Ser Glu Glu  
85 90 95  
Pro Pro Gly Ser Lys Gly Ser Val Thr Leu Ser Asp Leu Pro Gly Phe  
100 105 110  
Leu Gly Asp Leu Ala Ser Glu Glu Asp Ser Ile  
115 120

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2353 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGAAGGCGA TGTCACTATT GGAGAAGATG CACCAAATCT TTCTTTTAGC ACCAGTGTGG	60
GAAATGAGGA CGCCAGGACA GCCTGGCCCG AATTACAACA GAGCCATGCT GTTAATCAGC	120
TCAAAGATTT GTTGCGCCAA CAAGCAGATA AGGAAAGTGA AGTATCTCCG TCAAGAAGAA	180
GAAAAATGTC CCCCTTGAGG TCATTAGAAC ATGAGGAAAC CAATATGCCT ACTATGCACG	240
ACCTTGTTCA TACTATTAAAT GACCAGTCTC AATATATTCA TCATTTAGAG GCAGAAGTTA	300
AGTTCTGCAA GGAGGAATC TCTGGAATGA AAAATAAAAT ACAAGTAGTT GTGCTTGAAA	360
ACGAAGGGCT CCAGCAACAG CTAAATCTC AAAGACAAGA GGAGACACTG AGGGAACAAA	420
CACCTCTGGA TGCATCCGGA AACATGCACA ATTCTTGGAT TACAACAGGT GAAGATTCTG	480
GGGTGGGCGA AACCTCCAAA AGACCATTTT CCCATGACAA TGCAGATTTT GCCAAAGCTG	540
CATCTGCTGG TGAGCAGCTA GAACTGGAGA AGCTAAACT TACTTATGAG GAAAAGTGTG	600
AAATTGAGGA ATCCCAATTG AAGTTTTTGA GGAACGACTT AGCTGAATAT CAGAGAACTT	660
GTGAAGATCT TAAAGAGCAA CTAAAGCATA AAGAATTTCT TCTGGCTGCT AATACTTGTA	720
ACCGTGTGG TGGTCTTTGT TTGAAATGTG CTCAGCATGA AGCTGTCTT TCCCAAACCC	780
ATACTAATGT TCATATGCAG ACCATCGAAA GACTGGTTAA AGAAAGAGAT GACTTGATGT	840
CTGCACTAGT TTCCGTAAGG AGCAGCTTGG CAGATACGCA GCAAAGAGAA GCAAGTGCTT	900
ATGAACAGGT GAAACAAGTT TTGCAAATAT CTGAGGAAGC CAATTTTGAA AAAACCAAGG	960
CTTTAATCCA GTGTGACCAG TTGAGGAAGG AGCTGGAGAG GCAGGCGGAG CGACTTGAAA	1020
AAGAACTTGC ATCTCAGCAA GAGAAAAGGG CCATTGAGAA AGACATGATG AAAAAGGAAA	1080
TAACGAAAGA AAGGGAGTAC ATGGGATCAA AGATGTTGAT CTTGTCTCAG AATATTGCCC	1140
AACTGAGAGC CCAGGTGGAA AAGGTTACAA AGGAAAAGAT TTCAGCTATT AATCAACTGG	1200
AGGAAATTCA AAGCCAGCTG GCTTCTCGGG AAATGGATGT CACAAAGGTG TGTGGAGAAA	1260
TGCGCTATCA GCTGAATAAA ACCAACATGG AGAAGGATGA GGCAGAAAAG GAGCACAGAG	1320
AGTTCAGAGC AAAAATAAC AGGGATCTTG AAATTAAAGA TCAGGAAATA GAGAAATTGA	1380
GAATAGAACT GGATGAAAGC AAACAACACT TGAACAGGA GCAGCAGAAG GCAGCCCTGG	1440

```

CCAGAGAGGA GTGCCTGAGA CTAACAGAAC TGCTGGGCGA ATCTGAGCAC CAACTGCACC 1500
TCACCAGACA GGAAAAAGAT AGCATTCAGC AGAGCTTTAG CAAGGAAGCA AAGGCCCAAG 1560
CCCTTCAGGC CCAGCAAAGA GAGCAGGAGC TGACACAGAA GATACAGCAA ATGGAAGCCC 1620
AGCATGACAA AACTGAAAAT GAACAGTATT TGTTCGTGAC CTCCCAGAAT ACATTTTGA 1680
CAAAGTTAAA GGAAGAATGC TGTACATTAG CCAAGAACT GGAACAAATC TCTCAAAAAA 1740
CCAGATCTGA AATAGCTCAA CTCAGTCAAG AAAAAAGGTA TACATATGAT AAATTGGGAA 1800
AGTTACAGAG AAGAAATGAA GAATTGGAGG AACAGTGTGT CCAGCATGGG AGAGTACATG 1860
AGACGATGAA GCAAAGGCTA AGGCAGCTGG ATAAGCACAG CCAGGCCACA GCCCAGCAGC 1920
TGGTGCAGCT CCTCAGCAAG CAGAACCAGC TTCTCCTGGA GAGGCAGAGC CTGTCGGAAG 1980
AGGTGGACCG GCTGCGGACC CAGTTACCCA GCATGCCACA ATCTGATTGC TGACCTGGAT 2040
GGAACAGAGT GAAATAAATG ATTTACAAAG AGATATTTAC ATTCATCTGG TTTAGACTTA 2100
ATATGCCACA ACGCACCAG ACCTTCCCAG GGTGACACCG CCTCAGCCTG CAGTGGGGCT 2160
GGTCCTCATC AACCGGGGCG CTGTCCCCGC ACGCAGTCGG GCTGGAGCTG GAGTCTGACT 2220
CTAGCTGAGC AGAGCTCCTG GTGTATGTTT TCAGAAATGG CTGAAGTTA TGTGTTTAAA 2280
TCTGCTCATT CGTATGCTAG GTTATACATA TGATTTTCAA TAAATGAACT TTTTAAAGAA 2340
AAAAAAAAAA AAA 2353

```

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 615 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Ser Pro Leu Arg Ser Leu Glu His Glu Glu Thr Asn Met Pro Thr
1           5           10           15

Met His Asp Leu Val His Thr Ile Asn Asp Gln Ser Gln Tyr Ile His
          20           25           30

His Leu Glu Ala Glu Val Lys Phe Cys Lys Glu Glu Leu Ser Gly Met
          35           40           45

```

Lys Asn Lys Ile Gln Val Val Val Leu Glu Asn Glu Gly Leu Gln Gln  
 50 55 60  
 Gln Leu Lys Ser Gln Arg Gln Glu Glu Thr Leu Arg Glu Gln Thr Leu  
 65 70 75 80  
 Leu Asp Ala Ser Gly Asn Met His Asn Ser Trp Ile Thr Thr Gly Glu  
 85 90 95  
 Asp Ser Gly Val Gly Glu Thr Ser Lys Arg Pro Phe Ser His Asp Asn  
 100 105 110  
 Ala Asp Phe Gly Lys Ala Ala Ser Ala Gly Glu Gln Leu Glu Leu Glu  
 115 120 125  
 Lys Leu Lys Leu Thr Tyr Glu Glu Lys Cys Glu Ile Glu Glu Ser Gln  
 130 135 140  
 Leu Lys Phe Leu Arg Asn Asp Leu Ala Glu Tyr Gln Arg Thr Cys Glu  
 145 150 155 160  
 Asp Leu Lys Glu Gln Leu Lys His Lys Glu Phe Leu Leu Ala Ala Asn  
 165 170 175  
 Thr Cys Asn Arg Val Gly Gly Leu Cys Leu Lys Cys Ala Gln His Glu  
 180 185 190  
 Ala Val Leu Ser Gln Thr His Thr Asn Val His Met Gln Thr Ile Glu  
 195 200 205  
 Arg Leu Val Lys Glu Arg Asp Asp Leu Met Ser Ala Leu Val Ser Val  
 210 215 220  
 Arg Ser Ser Leu Ala Asp Thr Gln Gln Arg Glu Ala Ser Ala Tyr Glu  
 225 230 235 240  
 Gln Val Lys Gln Val Leu Gln Ile Ser Glu Glu Ala Asn Phe Glu Lys  
 245 250 255  
 Thr Lys Ala Leu Ile Gln Cys Asp Gln Leu Arg Lys Glu Leu Glu Arg  
 260 265 270  
 Gln Ala Glu Arg Leu Glu Lys Glu Leu Ala Ser Gln Gln Glu Lys Arg  
 275 280 285  
 Ala Ile Glu Lys Asp Met Met Lys Lys Glu Ile Thr Lys Glu Arg Glu  
 290 295 300  
 Tyr Met Gly Ser Lys Met Leu Ile Leu Ser Gln Asn Ile Ala Gln Leu  
 305 310 315 320  
 Glu Ala Gln Val Glu Lys Val Thr Lys Glu Lys Ile Ser Ala Ile Asn  
 325 330 335  
 Gln Leu Glu Glu Ile Gln Ser Gln Leu Ala Ser Arg Glu Met Asp Val



340	345	350
Thr Lys Val Cys Gly Glu Met Arg Tyr Gln Leu Asn Lys Thr Asn Met		
355	360	365
Glu Lys Asp Glu Ala Glu Lys Glu His Arg Glu Phe Arg Ala Lys Thr		
370	375	380
Asn Arg Asp Leu Glu Ile Lys Asp Gln Glu Ile Glu Lys Leu Arg Ile		
385	390	395
Glu Leu Asp Glu Ser Lys Gln His Leu Glu Gln Glu Gln Lys Ala		
405	410	415
Ala Leu Ala Arg Glu Glu Cys Leu Arg Leu Thr Glu Leu Leu Gly Glu		
420	425	430
Ser Glu His Gln Leu His Leu Thr Arg Gln Glu Lys Asp Ser Ile Gln		
435	440	445
Gln Ser Phe Ser Lys Glu Ala Lys Ala Gln Ala Leu Gln Ala Gln Gln		
450	455	460
Arg Glu Gln Glu Leu Thr Gln Lys Ile Gln Gln Met Glu Ala Gln His		
465	470	475
Asp Lys Thr Glu Asn Glu Gln Tyr Leu Leu Leu Thr Ser Gln Asn Thr		
485	490	495
Phe Leu Thr Lys Leu Lys Glu Glu Cys Cys Thr Leu Ala Lys Lys Leu		
500	505	510
Glu Gln Ile Ser Gln Lys Thr Arg Ser Glu Ile Ala Gln Leu Ser Gln		
515	520	525
Glu Lys Arg Tyr Thr Tyr Asp Lys Leu Gly Lys Leu Gln Arg Arg Asn		
530	535	540
Glu Glu Leu Glu Glu Gln Cys Val Gln His Gly Arg Val His Glu Thr		
545	550	555
Met Lys Gln Arg Leu Arg Gln Leu Asp Lys His Ser Gln Ala Thr Ala		
565	570	575
Gln Gln Leu Val Gln Leu Leu Ser Lys Gln Asn Gln Leu Leu Leu Glu		
580	585	590
Arg Gln Ser Leu Ser Glu Glu Val Asp Arg Leu Arg Thr Gln Leu Pro		
595	600	605
Ser Met Pro Gln Ser Asp Cys		
610	615	

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 313 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGACCTCTT CTGCGGCCGG CCTGGGCAGG TGTCTTCCTC GAGAGGCAGG CAGGGGATCC	60
CGGACACTAG CTTTATCGTC ATCTGGGAAA TTGTAAAAA TGCAAATTCG CAAGTTTGAG	120
AGCCATGGTT CCAAGAAACT GCATAAGCAT ACGAAATAAG TTGCAGCCTC CCGACTTATA	180
CCCTGGTACT TCTAGTCTAA AACAGGATTT GACTCTACTA ATCCAGCCTT ATACAGGATG	240
CTGTGTTCTT TGCTCCTTTG TGAATGTCTG TTGCTGGTAG CTGGTTATGC TCATGATGAT	300
GACTGGATTG ACC	313

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 677 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCTTGGATGA TGCATTAAGT GATATTTTAA TTAATTTTAA GTTTCATGAT TTTGAAACAT	60
GGAAGTGGCG ATTCAAGAT TCCTTTGGAG TGGATCCATA TAATGTGTTA ATGGTAATTC	120
TTTGTCTGCT CTGCATCGTG GTTTTAGTGG CTACTGAGCT GTGGACATAT GTATCTTGGT	180
ACACTCAGTT GAGACGTGTT TTAATCATCA GCTTCTGT T CAGTTTGGGA TGGAAATTGGA	240
TGTATTTATA TAAGCTAGCT TTTGCACAGC ATCAGGCTGA AGTCGCCAAG ATGGAGCCAT	300
TAAACAATGT GTGTGCCAAA AAGATGGACT GGACTGGAAG TATCTGGGAA TGTTTAGAA	360
GTTCATGGAC CTATAAGGAT GACCCATGCC AAAAATACTA TGAGCTCTTA CTAGTCAACC	420
CTATTTGGTT GGTCCCACCA ACAAAGGCAC TTGCAGTTAC ATTCAACCACA TTTGTAACGG	480

```

AGCCATTGAA GCATATTGGA AAAGGAACTG GGAATTTAT TAAAGCACTC ATGAAGGAAA      540
TTCCAGCGCT GCTTCATCTT CCAGTGCTGA TAATTATGGC ATTAGCCATC CTGAGTTTCT      600
GCTATGGTGC TGGAAAATCA GTTCATGTGC TGAGACATAT AGGCGGTCCT GAGAGCGAAC      660
CTCCCCAGGC ACTTCGG                                677

```

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 189 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

Met Val Ile Leu Cys Leu Leu Cys Ile Val Val Leu Val Ala Thr Glu
1              5              10              15
Leu Trp Thr Tyr Val Cys Trp Tyr Thr Gln Leu Arg Arg Val Leu Ile
                20              25              30
Ile Ser Phe Leu Phe Ser Leu Gly Trp Asn Trp Met Tyr Leu Tyr Lys
                35              40              45
Leu Ala Phe Ala Gln His Gln Ala Glu Val Ala Lys Met Glu Pro Leu
                50              55              60
Asn Asn Val Cys Ala Lys Lys Met Asp Trp Thr Gly Ser Ile Trp Glu
65              70              75              80
Trp Phe Arg Ser Ser Trp Thr Tyr Lys Asp Asp Pro Cys Gln Lys Tyr
                85              90              95
Tyr Glu Leu Leu Leu Val Asn Pro Ile Trp Leu Val Pro Pro Thr Lys
                100             105             110
Ala Leu Ala Val Thr Phe Thr Thr Phe Val Thr Glu Pro Leu Lys His
                115             120             125
Ile Gly Lys Gly Thr Gly Glu Phe Ile Lys Ala Leu Met Lys Glu Ile
                130             135             140
Pro Ala Leu Leu His Leu Pro Val Leu Ile Ile Met Ala Leu Ala Ile
145             150             155             160
Leu Ser Phe Cys Tyr Gly Ala Gly Lys Ser Val His Val Leu Arg His
                165             170             175

```

Ile Gly Gly Pro Glu Ser Glu Pro Pro Gln Ala Leu Arg  
 180 185

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

AGACGGCAGG AGGAATTGAT TATAGACCTG ATGGTGGAGC AGGTGATGCC GATTTCCATT      60
ATAGGGGCCA AATGGGCCCC ATTGAGCAAG GCCCTTATGC CAAAATGTAT GAGGGTAGAA      120
GAGAGATTTT GAGAGAGAGA GATGTTGACT TGAGATTTCA GGCTGGTCTC GAACTCCTGA      180
CCTCAAGTGA CCCGCCCTTG TCGGCCTCCC AAAGTGCTGG GATTACAGGC ATGAGCCATT      240
GTGCCCAGCC TATATAGTGT GAAGCTTTTA GGAAAATCAG AACAGGGTAG ACAGTTGTTA      300
AAAACAATGT TTAAATGGAA TAATGTTGAA TGTTTACAGG CTGTAAGAAT TATTGTATAC      360
ACAAAATAAT ACACAAAGTT TGTACTTTGT GTACAAATAC AAATTTGTAC TTTGTGTACA      420
AATAATACAA AAAGTTTGTA TACACAAAAA AAAAAAAAAA AAAAAAAAAA      470

```

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

CTGGAGTCCA CGCGGATTTT CGAAGCTGGG GCTGGCAAGA GGCCGCTGGA CACCACGCTC      60
CAGTCGTCAG CCCACTTCCT AGCTGAACAG CGCGAGGCGG CGGCAGCGAG CCGGGTCCCA      120
CCATGGCCGC GAATTATTCC AGTACCAGTA CCCGGAGAGA ACATGTCAAA GTTAAACCA      180

```

GCTCCCAGCC AGGCTTCCTG GAACGGCTGA GCGAGACCTC GGGTGGGATG TTTGTGGGGC 240  
TCATGGCCTT CCTGCTCTCC TTCTACCTAA TTTCACCAA TGAGGGCCGC GCATTGAAGA 300  
CGGCAACCTC ATTGGCTGAG GGGCTCTCGC TTGTGGTGTG TCCTGACAGC ATCCACAGTG 360  
TGGCTCCGGA GAATGAAGGA AGGCTGGTGC ACATCATTGG CGCCTTACGG ACATCCAAGC 420  
TTTGTCTGA TCCAAACTAT GGGGTCCATC TTCCGGCTGT GAAACTGCGG AGGCACGTGG 480  
AGATGTACCA ATGGGTAGAA ACTGAGGAGT CCAGGGAGTA CACCGAGGAT GGGCAGGTGA 540  
AGAAGGAGAC GAGGTATTCC TACAACACTG AATGGAGGTC AGAAATCATC AACAGCAAAA 600  
ACTTCGACCG AGAGATTGGC CACAAAAACC CCAGCTTCCT CTCTCCCACA GTGCCATGGC 660  
AGTGGAGTCA TTCATGGCAA CAGCCCCCTT TGTCCAAAT TGGCAGGTTT TCCTCTCGTC 720  
AGGCCTCATC GACAAAGTCG ACAACTTCAA GTCCCTGAGC CTATCCAAGC TGGAGGACCC 780  
TCATGTGGAC ATCATTCGCC GTGGAGACTT TTTCTACCAC AGCGAAAATC CCAAGTATCC 840  
AGAGGTGGGA GACTTGCCTG TCTCCTTTTC CTATGCTGGA CTGAGCGCG ATGACCCTGA 900  
CCTGGGCCCA GCTCACGTGG TAACCTGGCT TCCCAGGGGC AGACACTAAG TCAGAGCCTC 960  
ACGACTTTCC TGGACACAGA CACCTTGGTC AATGTCAGGA GCGCTTGGAC CCCCTTTTCC 1020  
CTGGGGAAAG GCACACTCTC GCACACACTC TCAGCCAGGC ACGCTTCTGA GCAGTTTCAG 1080  
AGCTCCCATG TCCCCACAGC CATCCATGGA CCCACGTTA AGAAGGGCAG CTCAAAGGG 1140  
GTCTCATAGT CGCACCTTAT GACAGGTGTT CCAGTCACAC ACAGACCCTC TCCCCAAGCC 1200  
CGTTTTGATC TGTCAATAAT TGGTCTTGCG TTCTTGGCCT ATGTGCAGTC CTGCCCCATC 1260  
CCCTGCTCTG CGCACTGCCC AAGAGCTTGG AATGCCTGGA GCTTTGAATG GAGCAGCTCA 1320  
GCCAGAGCTG CAGAGGTGGA TGCATCCCAG ATGGATGTAT AGAGAGAGAA GCCCCAGGGT 1380  
CTCTGTGCTC ACTTCCCCAG CCGGCACCCA GTCCCGGGAG GGTGGGCCAT GGCTCTCATG 1440  
GGCGTGTCTC CCGCTGGTCA CCCCTCAGCT CTAACACCAG GTCTCTGAC CAGGTCACTG 1500  
TGATTGCCCC GCAGCGGGGT GACCAGCTAG TCCCATTCTC CACCAAGTCT GGGGATACCT 1560  
TACTGCTCCT GCACCACGGG GACTTCTCAG CAGAGGAGGT GTTTCATAGA GAACTAAGGA 1620  
GCAACTCCAT GAAGACCTGG GGCCTGCGGG CAGCTGGCTG GATGGCCATG TTCATGGGCC 1680  
TCAACCTTAT GACACGGATC CTCTACACCT TGGTGGACTG GTTTCCTGTT TTCCGAGACC 1740  
TGGTCAACAT TGGCCTGAAA GCCTTTGCCT TCTGTGTGGC CACCTCGCTG ACCCTGCTGA 1800  
CCGTGGCGGC TGGCTGGCTC TTCTACCGAC CCCTGTGGGC CCTCCTCATT GCCGGCCTGG 1860

```

CCCTTGTGCC CATCCTTGTT GCTCGGACAC GGGTGCCAGC CAAAAAGTTG GAGTGAAAAG 1920
ACCCTGGCAC CCGCCCGACA CCTGCGTGAG CCCTAGGATC CAGGTCCTCT CTCACCTCTG 1980
ACCCAGCTCC ATGCCAGAGC AGGAGCCCCG GTCAATTTTG GACTCTGCAC CCCCTCTCCT 2040
CTTCAGGGGC CAGACTTGGC AGCATGTGCA CCAGGTTGGT GTTCACCAGC TCATGTCTTC 2100
CCCACATCTC TTCTTGCCAG TAAGCAGCTT TGGTGGGCAG CAGCAGCTCA TGAATGGCAA 2160
GCTGACAGCT TCTCCTGCTG TTTCCTTCCT CTCTTGGACT GAGTGGGTAC GGCCAGCCAC 2220
TCAGCCCATT GGCAGCTGAC AACGCAGACA CGCTCTACGG AGGCCTGCTG ATAAAGGGCT 2280
CAGCCTTGCC GTGTGCTGCT TCTCATCACT GCACACAAGT GCCATGCTTT GCCACCACCA 2340
CCAAGCACAT CTGTGATCCT GAAGGGCGGC CGTTAGTCAT TACTGCTGAG TCCTGGGTCA 2400
CCAGCAGACA CACTGGGCAT GGACCCCTCA AAGCAGGCAC ACCCAAAACA CAAGTCTGTG 2460
GCTAGAACCT GATGTGGTGT TTAAGAGAGA AGAAACACTG AAGATGTCCT GAGGAGAAAA 2520
GCTGGACATA TACTGGGCTT CACACTTATC TTATGGCTTG GCAGAATCTT TGTAGTGTGT 2580
GGGATCTCTG AAGGCCCTAT TTAAGTTTTT CTTCGTTACT TTGCTGCTTC ATGTGTACTT 2640
TCCTACCCCA AGAGGAAGTT TTCTGAAATA AGATTTAAAA ACAAACAAA AAAAAAAAAA 2700
AA 2702

```

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Ala Ala Asn Tyr Ser Ser Thr Ser Thr Arg Arg Glu His Val Lys
1           5           10           15
Val Lys Thr Ser Ser Gln Pro Gly Phe Leu Glu Arg Leu Ser Glu Thr
20           25           30
Ser Gly Gly Met Phe Val Gly Leu Met Ala Phe Leu Leu Ser Phe Tyr
35           40           45
Leu Ile Phe Thr Asn Glu Gly Arg Ala Leu Lys Thr Ala Thr Ser Leu

```

50	55	60
Ala Glu Gly Leu Ser Leu Val Val Ser Pro Asp Ser Ile His Ser Val		
65	70	75 80
Ala Pro Glu Asn Glu Gly Arg Leu Val His Ile Ile Gly Ala Leu Arg		
	85	90 95
Thr Ser Lys Leu Leu Ser Asp Pro Asn Tyr Gly Val His Leu Pro Ala		
	100	105 110
Val Lys Leu Arg Arg His Val Glu Met Tyr Gln Trp Val Glu Thr Glu		
	115	120 125
Glu Ser Arg Glu Tyr Thr Glu Asp Gly Gln Val Lys Lys Glu Thr Arg		
	130	135 140
Tyr Ser Tyr Asn Thr Glu Trp Arg Ser Glu Ile Ile Asn Ser Lys Asn		
	145	150 155 160
Phe Asp Arg Glu Ile Gly His Lys Asn Pro Ser Phe Leu Ser Pro Thr		
	165	170 175
Val Pro Trp Gln Trp Ser His Ser Trp Gln Gln Pro Pro Leu Ser Lys		
	180	185 190
Leu Ala Gly Phe Ser Ser Arg Gln Ala Ser Ser Thr Lys Ser Thr Thr		
	195	200 205
Ser Ser Pro		
210		

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3395 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATCTTCCTGC CCTTCACCTG CATTTGGCTAC ACGGCCACCA ATCAGGACTT CATCCAGCGC	60
CTGAGCACAC TGATCCGGCA GGCCATCGAG CGGCAGCTGC CTGCCTGGAT CGAGGCTGCC	120
AACCAGCGGG AGGAGGGCCA GGGTGAACAG GGCGAGGAGG AGGATGAGGA GGAGGAAGAA	180
GAGGAGGACG TGGCTGAGAA CCGCTACTTT GAAATGGGGC CCCCAGACGT GGAGGAGGAG	240

GAGGGAGGAG GCCAGGGGGA GGAAGAGGAG GAGGAAGAGG ARGATGAAGA RGCCGAGGAG	300
GAGCGCCTGG CTCTGGAATG GGCCCTGGGC GCGGACGAAG ACTTCCTGCT GGAGCACATC	360
CGCATCCTCA AGGTGCTGTG GTGCTTCCTG ATCCATGTGC AGGGCAGTAT CCGCCAGTTC	420
GCCGCTGCC TTGTGCTCAC CGACTTCGGC ATCGCAGTCT TCGAGATCCC GCACCAGGAG	480
TCTCGGGGCA GCAGCCAGCA CATCCTCTCC TCCCTGCGCT TTGTCTTTTG CTTCCTCGAT	540
GGCGACCTCA CCGAGTTTGG CTTCTTCATG CCGGAGCTGT GTCTGGTGCT CAAGGTACGG	600
CACAGTGAGA ACACGCTCTT CATATCTCG GACGCCGCCA ACCTGCACGA GTTCCACGCG	660
GACCTGCGCT CATGCTTTGC ACCCCAGCAC ATGGCCATGC TGTGTAGCCC CATCCTCTAC	720
GGCAGCCACA CCAGCCTGCA GGAGTTCCTG CGCCAGCTGC TCACCTTCTA CAAGGTGGCT	780
GGCGGCTGCC AGGAGCGCAG CCAGGGCTGC TCCCCGTCT ACCTGGTCTA CAGTGACAAG	840
CGCATGGTGC AGACGGCCGC CGGGGACTAC TCAGGCAACA TCGAGTGGGC CAGCTGCACA	900
CTCTGTTTCA CCGTGCGGCG CTCCTGCTGC GCGCCCTCTG AGGCCGTCAA GTCCGCCGCC	960
ATCCCCTACT GGCTGTTGCT CACGCCCCAG CACCTCAACG TCATCAAGGC CGACTTCAAC	1020
CCCATGCCCA ACCGTGGCAC CCACAACGTG CGCAACCGCA ACAGCTTCAA GCTCAGCCGT	1080
GTGCCGCTCT CCACCGTGCT GCTGGACCCC ACACGCAGCT GTACCCAGCC TCGGGGCGCC	1140
TTTGCTGATG GCCACGTGCT AGAGCTGCTC GTGGGGTACC GCTTTGTAC TGCCATCTTC	1200
GTGCTGCCCC ACGAGAAGTT CCACTTCCTG CGCGTCTACA ACCAGCTGCG GGCTCGCTG	1260
CAGGACCTGA AGACTGTGGT CATCGCCAAG ACCCCCGGGA CGGGAGGCAG CCCCAGGGC	1320
TCCTTTGCGG ATGGCCAGCC TGCCGAGCGC AGGGCCAGCA ATGACCAGCG TCCCAGGAG	1380
GTCCCAGCAG AGGCTCTGGC CCCGGCCCCA GTGGAAGTCC CAGCTCCAGC CCCTGCAGCA	1440
GCCTCAGCCT CAGGCCCAGC GAAGACTCCG GCCCCAGCAG AGGCCTCAAC TTCAGCTTTG	1500
GTCCCAGAGG AGACGCCAGT GGAAGCTCCA GCCCCACCC CAGCCGAGGC CCCTGCCCAG	1560
TACCCGAGTG AGCACCTCAT CCAGGCCACC TCGGAGGAGA ATCAGATCCC CTCGCACTTG	1620
CCTGCCTGCC CGTCGCTCCG GCACGTCGCC AGCCTGCGGG GCAGCGCCAT CATCGAGCTC	1680
TTCCACAGCA GCATTGCTGA GGTGAAAAC GAGGAGCTGA GGCACCTCAT GTGGTCCTCG	1740
GTGGTGTTCT ACCAGACCCC AGGGCTGGAG GTGACTGCCT GCGTGCTGCT CTCCACCAAG	1800
GCTGTGTACT TTGTGCTCCA CGACGGCCTC CGCCGCTACT TCTCAGAGCC ACTGCAGGAT	1860
TTCTGGMATC AGAAAAACAC SGAATAAAC AACAGCCCTT TCCACATCTC CCAGTGCTTC	1920



GTGCTAAAGC TTAGTGACCT GCAGTCAGTC AATGTGGGGC TTTTCGACCA GCATTTCGG	1980
CTGACGGGTT CCACCCGAT GCAGGTGGTM ACGTGC TTGA CGCGGGACAG CTACCTGACG	2040
CACTGCTTCC TCCAGCACCT CATGGTCGTG CTGTCTCTC TGGAAACGCAC GCCCTCGCCG	2100
GAGCCTGTTG ACAAGGACTT CTA CTCCGAG TTTGGGAACA AGACCACAGG GAAGATGGAG	2160
AACTACGAGC TGATCCACTC TAGTCGCGTC AAGTTTACCT ACCCCAGTGA GGAGGAGATT	2220
GGGACCTGA CGTTCACTGT GGCCCAAAAG ATGGCTGAGC CAGAGAAGGC CCCAGCCCTC	2280
AGCATCTGCG TGACGTGCA GGCCTTCCAG GTGGGCATGC CACCCCTGG GTGCTGCAGG	2340
GGCCCCCTGC GCCCAAGAC ACTCTGCTC ACCAGCTCCG AGATCTTCTT CCTGGATGAG	2400
GA CTGTGTCC ACTACCCACT GCCCGAGTTT GCCAAAGAGC CGCCGCAGAG AGACAGGTAC	2460
CGGCTGGACG ATGGCCGCCG CGTCCGGGAC CTGGACCGAG TGCTCATGGG CTACCAGACC	2520
TACCCGCAGG CCTTCACCCT CGTCTTCGAT GACGTGCAAG GTCATGACCT CATGGGCAGT	2580
GTCACCCTGG ACCACTTTGG GGAGGTGCCA GGTGGCCCG CTAGAGCCAG CCAGGGCCGT	2640
GAAGTCCAGT GGCAGGTGT TGTCCCCAGT GCTGAGAGCA GAGAGAAGCT CATCTCGCTG	2700
TTGGCTCGCC AGTGGGAGGC CCTGTGTGGC CGTGAGCTGC CTGTCGAGCT CACCGCTAG	2760
CCCAGGCCAC AGCCAGCTG TCGTGTCCAG CCTGACGCCT ACTGGGGCAG GGCAGCAGGC	2820
TTTTGTGTTC TCTAAAAATG TTTTATCCTC CCTTTGGTAC CTTAATTTGA CTGTCTCGC	2880
AGAGAAATGTG AACATGTGTG TGTGTGTGT TAATTCTTTC TCATGTTGGG AGTGAGAATG	2940
CCGGGCCCCCT CAGGGCTGTC GGTGTGCTGT CAGCCTCCCA CAGGTGGTAC AGCCGTGCAC	3000
ACCAGTGTG TGTCTGCTGT TGTGGGACCG TTGTTAACAC GTGACACTGT GGGTCTGACT	3060
TTCTCTTCTA CACGTCTTT CCTGAAGTGT CGAGTCCAGT CCTTTGTTGC TGTGCTGTT	3120
GCTGTGCTG TTGCTGTTGG CATCTGCTG CTAATCCTGA GGCTGGTAGC AGAATGCACA	3180
TTGGAAGCTC CCACCCATA TTGTCTTCA AAGTGGAGGT CTCCCCTGAT CCAGACAAGT	3240
GGGAGAGCCC GTGGGGGCAG GGGACCTGGA GCTGCCAGCA CCAAGCGTGA TTCCTGCTGC	3300
CTGTATTCTC TATTCCAATA AAGCAGAGTT TGACACCGTC AAAAAAAAAA AAAAAAAAAA	3360
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA	3395

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 848 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Gly Pro Pro Asp Val Glu Glu Glu Glu Gly Gly Gly Gln Gly Glu
1           5           10           15

Glu Glu Glu Glu Glu Glu Glu Asp Glu Glu Ala Glu Glu Glu Arg Leu
20           25           30

Ala Leu Glu Trp Ala Leu Gly Ala Asp Glu Asp Phe Leu Leu Glu His
35           40           45

Ile Arg Ile Leu Lys Val Leu Trp Cys Phe Leu Ile His Val Gln Gly
50           55           60

Ser Ile Arg Gln Phe Ala Ala Cys Leu Val Leu Thr Asp Phe Gly Ile
65           70           75           80

Ala Val Phe Glu Ile Pro His Gln Glu Ser Arg Gly Ser Ser Gln His
85           90           95

Ile Leu Ser Ser Leu Arg Phe Val Phe Cys Phe Pro His Gly Asp Leu
100          105          110

Thr Glu Phe Gly Phe Leu Met Pro Glu Leu Cys Leu Val Leu Lys Val
115          120          125

Arg His Ser Glu Asn Thr Leu Phe Ile Ile Ser Asp Ala Ala Asn Leu
130          135          140

His Glu Phe His Ala Asp Leu Arg Ser Cys Phe Ala Pro Gln His Met
145          150          155          160

Ala Met Leu Cys Ser Pro Ile Leu Tyr Gly Ser His Thr Ser Leu Gln
165          170          175

Glu Phe Leu Arg Gln Leu Leu Thr Phe Tyr Lys Val Ala Gly Gly Cys
180          185          190

Gln Glu Arg Ser Gln Gly Cys Phe Pro Val Tyr Leu Val Tyr Ser Asp
195          200          205

Lys Arg Met Val Gln Thr Ala Ala Gly Asp Tyr Ser Gly Asn Ile Glu
210          215          220

Trp Ala Ser Cys Thr Leu Cys Ser Ala Val Arg Arg Ser Cys Cys Ala
225          230          235          240

```

Pro Ser Glu Ala Val Lys Ser Ala Ala Ile Pro Tyr Trp Leu Leu Leu  
 245 250 255  
 Thr Pro Gln His Leu Asn Val Ile Lys Ala Asp Phe Asn Pro Met Pro  
 260 265 270  
 Asn Arg Gly Thr His Asn Cys Arg Asn Arg Asn Ser Phe Lys Leu Ser  
 275 280 285  
 Arg Val Pro Leu Ser Thr Val Leu Leu Asp Pro Thr Arg Ser Cys Thr  
 290 295 300  
 Gln Pro Arg Gly Ala Phe Ala Asp Gly His Val Leu Glu Leu Leu Val  
 305 310 315 320  
 Gly Tyr Arg Phe Val Thr Ala Ile Phe Val Leu Pro His Glu Lys Phe  
 325 330 335  
 His Phe Leu Arg Val Tyr Asn Gln Leu Arg Ala Ser Leu Gln Asp Leu  
 340 345 350  
 Lys Thr Val Val Ile Ala Lys Thr Pro Gly Thr Gly Gly Ser Pro Gln  
 355 360 365  
 Gly Ser Phe Ala Asp Gly Gln Pro Ala Glu Arg Arg Ala Ser Asn Asp  
 370 375 380  
 Gln Arg Pro Gln Glu Val Pro Ala Glu Ala Leu Ala Pro Ala Pro Val  
 385 390 395 400  
 Glu Val Pro Ala Pro Ala Pro Ala Ala Ala Ser Ala Ser Gly Pro Ala  
 405 410 415  
 Lys Thr Pro Ala Pro Ala Glu Ala Ser Thr Ser Ala Leu Val Pro Glu  
 420 425 430  
 Glu Thr Pro Val Glu Ala Pro Ala Pro Pro Pro Ala Glu Ala Pro Ala  
 435 440 445  
 Gln Tyr Pro Ser Glu His Leu Ile Gln Ala Thr Ser Glu Glu Asn Gln  
 450 455 460  
 Ile Pro Ser His Leu Pro Ala Cys Pro Ser Leu Arg His Val Ala Ser  
 465 470 475 480  
 Leu Arg Gly Ser Ala Ile Ile Glu Leu Phe His Ser Ser Ile Ala Glu  
 485 490 495  
 Val Glu Asn Glu Glu Leu Arg His Leu Met Trp Ser Ser Val Val Phe  
 500 505 510  
 Tyr Gln Thr Pro Gly Leu Glu Val Thr Ala Cys Val Leu Leu Ser Thr  
 515 520 525  
 Lys Ala Val Tyr Phe Val Leu His Asp Gly Leu Arg Arg Tyr Phe Ser

530	535	540
Glu Pro Leu Gln Asp Phe Trp Xaa Gln Lys Asn Thr Asp Tyr Asn Asn 545	550	555 560
Ser Pro Phe His Ile Ser Gln Cys Phe Val Leu Lys Leu Ser Asp Leu 565	570	575
Gln Ser Val Asn Val Gly Leu Phe Asp Gln His Phe Arg Leu Thr Gly 580	585	590
Ser Thr Pro Met Gln Val Val Thr Cys Leu Thr Arg Asp Ser Tyr Leu 595	600	605
Thr His Cys Phe Leu Gln His Leu Met Val Val Leu Ser Ser Leu Glu 610	615	620
Arg Thr Pro Ser Pro Glu Pro Val Asp Lys Asp Phe Tyr Ser Glu Phe 625	630	635 640
Gly Asn Lys Thr Thr Gly Lys Met Glu Asn Tyr Glu Leu Ile His Ser 645	650	655
Ser Arg Val Lys Phe Thr Tyr Pro Ser Glu Glu Glu Ile Gly Asp Leu 660	665	670
Thr Phe Thr Val Ala Gln Lys Met Ala Glu Pro Glu Lys Ala Pro Ala 675	680	685
Leu Ser Ile Leu Leu Tyr Val Gln Ala Phe Gln Val Gly Met Pro Pro 690	695	700
Pro Gly Cys Cys Arg Gly Pro Leu Arg Pro Lys Thr Leu Leu Leu Thr 705	710	715 720
Ser Ser Glu Ile Phe Leu Leu Asp Glu Asp Cys Val His Tyr Pro Leu 725	730	735
Pro Glu Phe Ala Lys Glu Pro Pro Gln Arg Asp Arg Tyr Arg Leu Asp 740	745	750
Asp Gly Arg Arg Val Arg Asp Leu Asp Arg Val Leu Met Gly Tyr Gln 755	760	765
Thr Tyr Pro Gln Ala Leu Thr Leu Val Phe Asp Asp Val Gln Gly His 770	775	780
Asp Leu Met Gly Ser Val Thr Leu Asp His Phe Gly Glu Val Pro Gly 785	790	795 800
Gly Pro Ala Arg Ala Ser Gln Gly Arg Glu Val Gln Trp Gln Val Phe 805	810	815
Val Pro Ser Ala Glu Ser Arg Glu Lys Leu Ile Ser Leu Leu Ala Arg 820	825	830

Gln Trp Glu Ala Leu Cys Gly Arg Glu Leu Pro Val Glu Leu Thr Gly  
 835 840 845

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:.

- (A) LENGTH: 1147 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

GGAAGGAGTT CTGGAATTGG AAAACCATCA TTTTTCACCC ATCACAGTAA ATATGGCTCA      60
GCCAAGAATT ATCAATCAAT GCTAAAGCTA GGGGGAAATT TCGCTTAGGA GCAGGATATT      120
AGGGTATTAG TCTGGGCTTA AACTATCTCC TCACAGATTG TTGTTAGTTT CTGGGGAAAG      180
AATAGTAACC ATGCAATGGA AAAAAATGGA CAACCTCTTG ACTAGGTTAT CAAAATTAAC      240
CTCACCAATA AAGGGTGGAT GTTCAACATG TGCCTTCAAA TGTGACCCAC TGAGAAGGAA      300
ACAACATCAC TGTAACAACA ACAACCAGAA ACGACAGGGG GTTTTGACTG AATTCCTTCAA      360
AAATGTCAAT GTCATAGAAG ACAAGAAAG GTTGTGGAAA TGTTCAGAT TAAATGATAG      420
TAAAAACACC TGACAACTAA ACATAGTAAG TAATACTAGA CTGGATTCTG TACCAGAGGT      480
AACATAAGTG CTCCAAAGGA CAATGTTAGG TCAACTGGCA AATTGGAATA TAGACAGTCA      540
ATCAGATAAG AAGTATACTT TGATTAAGTA AAAAAAATCC CTATTCTTGG AAAATACACA      600
ATAAAGTATT TTGAGGTAAA GGGCCATAAT GTATGCAATC TACTCTCAA AAATTCAGAA      660
ACATATATTT GTGTGCATTT GCATGTGCAA CAGTACACAC AAACATACAT AAAGAGAGCA      720
ATTGATAAGG CAAATAAGGT AACATTTAAC AATAATCTGA TACACATAAA TAGAGAAAGA      780
GCAATTGATA AAGTAAATGA GGTAAAATTT AACAATAATC TGAGCAAAAG GTATATGTGT      840
TTTCTTTGAG ACAGTCTGAT TCTTGCAACT TATTCTGTAA GTTGGAACTT ATTTCCAAAC      900
ATGATTGAAA AAAAACCCCG CACTTGGCAA CTTCCTCTCT TTTTCAGCCT AGAAATGTCT      960
GTGTTAAGTG GTTTTTTATT TATTGTTGTT GTTTGTTGTT ATGTTGTTT TGTGCCAGG      1020
CTCCAATCA CAAAATACGA GTTAAAAAAC TGCCTTGTTA TTTTATAGAGA TTTGTGATAA      1080

```

TACAACTTGT TATAAAATTT ATTCTCAAT AAATATAATT TCTCTACTAT GCAAAAAAAA 1140  
 AAAAAAA 1147

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ile Glu Lys Lys Pro Arg Thr Trp Gln Leu Leu Phe Phe Ser  
 1 5 10 15  
 Leu Glu Met Ser Val Leu Ser Gly Phe Leu Phe Ile Val Val Val Cys  
 20 25 30  
 Cys Tyr Cys Cys Phe Val Ala Arg Leu Gln Leu Thr Lys Tyr Glu Phe  
 35 40 45  
 Lys Asn Cys Val Val Ile Phe Arg Asp Leu  
 50 55

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1013 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCTTTTAAAA AAATATCTGA AAAAAGCTTC ATATCTTTAC AAATCATAA AATAGCTGAT 60  
 TGGGCCATGG AGGAGATGAG GCTGTTTAGA ACTGGTTTIG TTTCAAGTTT GTCAATTTTC 120  
 CCTGTATGAG AACTTGGGTA AAGCACAAAG AAACATACAG TGCTAGTAAC AGGTCTCCTG 180  
 CGCCCTGGAA CTAAGTGTTC GGAGGAAGGA CTAAACCCCG GGGGAGGTGA GTATAAATA 240  
 ATTCCACTAA GATCACCTCC TCAGTCCCCA GAAGGCTGAT GGTGGATCCT CTGGCCATCT 300

```

CCTGTGGGGT CTTACTGCTC CTCTGCCATT TCTCTATGCC TGAAGACACG AAGATGATAT   360
CAAGGCAGAG CTACCATATC GCAGCCAGTC TCTAGGCTAC TGCTGTGCAG TGGCTCCAC   420
TTTCTAATGC TTTTTTGTTC TTGCTTTTTC TAACAAAACA ATCTTTTTC AAAATGAATT   480
CCAACCCCTG CTAGTTCCTT CGCTGCCTCC ATACTGTTTT AGGCAGCACC GTTTATGTGA   540
CAGAGTCCGT GTTTCTCAAA TGCATGGTGT TCCTCAGGTG GAGAGTGGGC AGAAGTTTTT   600
GCAACACTTT TTTTTTAAGT TATTGGGTGC AAAATCCCAA ACCAGGATAT GTGTATGTCT   660
GTGTGTTTAT GTTTTTTATT TGACCTCCC CTCTTCAAC CTACCCCTT TTATATCTAA   720
TGTAGAAAAA GCGAAATTGA ATCTGGAAAG CAACTGTTG TATATAGTTG CGGTAACAAT   780
CATGAAGAGA GAGCCGGGCT GTCCCTCAG TAATTCATTT TAAATAACAA ATTATTTAAA   840
AATAAAATTC ATGCCAGAGC CAGCTGAAGA GGCCTTCCTT CATCACCCT GAGGCCACCC   900
CCAATCTGGG CCCTCTGTCC ATCTGGCATG TCTCCTCCA GCAAGATTCA TCTGTTCAAT   960
GCCATTTGCG TTTCAATAAA GTTATCTCCT GTACTGTCAA AAAAAAAAAA AAA   1013

```

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met His Gly Val Pro Gln Val Glu Ser Gly Gln Lys Phe Leu Gln His
1           5           10           15

Phe Phe Phe Lys Leu Leu Gly Ala Lys Ser Gln Thr Arg Ile Cys Val
20           25           30

Cys Leu Cys Val Tyr Val Phe Tyr Leu Thr Leu Pro Ser Phe Asn Leu
35           40           45

Pro Pro Phe Ile Ser Asn Val Glu Lys Ala Lys Leu Asn Leu Glu Ser
50           55           60

Lys Leu Leu Tyr Ile Val Ala Val Thr Ile Met Lys Arg Glu Pro Gly
65           70           75           80

Cys Pro Leu Ser Asn Ser Phe

```

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1763 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

TCGGGATAAA AAGCAAGAAA AGAAAGAGAA GACTGAGAAT AAGAAGATCT CTTTGAAAAAT      60
AAAATAAGAC TGCTAAAAGT ATTTGGTATA CAGTCTGGAA AATAAAGTTG AGGGAATCTC      120
TCCAGATAAA GAGCAAAAAG AAATAGATAG AAAAATATAA AGAAAGAAAA AAGACATAGA      180
CAATCAATAT GTAATGTTAG GAGTTCCTGG AAGAGAGAAC AGAGACAGTG TAGGTGAAGA      240
AATAAAAAGA AAAAGAATTG AAGAACAGAG CAAGCTAAGT CTCCAGATTG AGAGGGCCCA      300
ATACAATCTA CATCTAGACA CAATATTGTA AAATTTTGGA ATATTAAGGA TAGAAGGAAG      360
ATATTAAAGT GGCCAGGGAG AAAACAAATG AGGTCATCAC GATTAGCTCA ACACAAAAAT      420
GGATGAGAAA TAGACTGCTA ACAGATTTGT CATCAGCAAC ACTGAATGCC AGAAGTCAAT      480
GGATCAACAT CTTCAGAGCT TAAGGAAAAT TTTTGTACCT AGAATTTTCA AGTAAGGCAG      540
ACTGTCAAGA AGAACATCAA AGTGAAGACA TTTTCTGTCA GGCAAATTTT CAGAAAGTCT      600
CCTTTGCACC CTTACTGAGG AAGTATCTTG AGGAAATTCT CCAGCAAAAT GAGGATGAAA      660
ACCAGGAAAG AAGAAGAAAT GGGATCCATA AAACAGTGGA CCTTACTTAG GATGTCTCAT      720
TCTAGAGTGA CAGCCAAAAG GGTATCTCAC CCTAGAGTGA CAGCTATCCA GCAGACTAAT      780
TTCAGATGAG AGCATACTGT CTCGGGCTTT CTGGAAGAA TGTGCATTCA GTGCCATAGA      840
TAGTATCACT GAAGAGCTGG GATGCTTGAG AAGATTATTT AGTCAAGAAA AAAGAAAGAC      900
AAATCAACAA TATGTCAAAA AATTCAGGTC CAATTATAGA GCAAAATAAA ATGAGGCATG      960
ATTTTGAGTT ATTATGAAG AATAAGAAGA GGCTTGATAG GTACATTTCC TTTTCTATGG     1020
CACAGGCATG ATGATATTGG GTGTGTAGGG AAGAAAATAT CCTAGCTTAT ACTAGGCTCC     1080
CAGTAAGAAG TATTTAAATA GCCAAAATAA TGTGGATATC ATTTATTAGT ATTCAATGTT     1140

```



CAGATCAGCC TATTAACAAA GTGTGAAAGG TTTCATTTTT TATTCAGAAC TGAAGTTGAA 1200  
 AGTAATTAAAT GCTGACAAAG GGAAAGAAAG CAGAAAGAGA TTGAGAATTA GAGGAAGAGA 1260  
 AGTGAATCA AAGGTAGAGA TACTTATATA TTCAAAGTGG GGATGAAAAG ATCTTCAGTT 1320  
 AATGAACAA GAAC TAGAGG ATTAGTGTAT TGTTCAAAGC TATAAAATCA AACCAATAGA 1380  
 TGTATTAAAA AGTGATGTAA CTATCAGACA TTTGGAGAGA GATGGACAAA GGAAAGTGCC 1440  
 GATAGTGTAA GTTAAATCCT TATCTTTTGT AATGGGGAAT TATTAAAGAT GTTGTAAGT 1500  
 CAGTAAGTCA AGAAATTATT GCTCAAACAT ATTATTTTAA GTTAGAAAGT TACCAGACGA 1560  
 TCTAAAATAA ATATTGTTAA AAGCATTACC TCTAGGGAAT GGGATTTAGA TTTAAAAAGG 1620  
 GTGGGATGGG AAACGTGTT TTTCATTTTA AGTCCTTCTG TACTATTTAA TTTTTTACCT 1680  
 TGTGCATGTA TTAACTTGAA AAAATTTTAA ATAAACCCAA ATAAAAATCT AAAAAAAAAA 1740  
 AAAAAAAAAA AAAAAAAAAA AAA 1763

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Arg Met Lys Thr Arg Lys Glu Glu Glu Met Gly Ser Ile Lys Gln  
 1 5 10 15  
 Trp Thr Leu Leu Arg Met Ser His Ser Arg Val Thr Ala Lys Arg Val  
 20 25 30  
 Ser His Pro Arg Val Thr Ala Ile Gln Gln Thr Asn Phe Arg  
 35 40 45

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNATCTTAGAG CTCAAAGTTA GGGTCTG

28

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNCAGAGCTGT TCTGATACTA AGTCTCAC

29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANACTATCTTC TTCAGAGGCC AGATCACC

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CNAGAAGCCAG CTGGCTTTGA ATTCCTC

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CNTTTTCCAAT ATGCTTCAAT GGCTCCGT

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNGGTAGAAGG AGAGCAGGAA GGCCATGA

29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNCTTCTCTGG CTCAGCCATC TTTTGGGC

29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CNGTACACACA AACATACATA AAGAGAGC

29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ANACGGACTCT GTCACATAAA CGGTGCTG

29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GNTGAGATACC CTTTGGCTG TCACTCTA

29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met	Gln	Trp	Lys	Lys	Met	Asp	Asn	Leu	Leu	Thr	Arg	Leu	Ser	Lys	Leu	1	5	10	15
Thr	Ser	Pro	Ile	Lys	Gly	Gly	Cys	Ser	Thr	Cys	Ala	Phe	Lys	Cys	Asp	20	25	30	
Pro	Leu	Arg	Arg	Lys	Gln	His	His	Cys	Asn	Asn	Asn	Asn	Gln	Lys	Arg	35	40	45	
Gln	Gly	Val	Leu	Thr	Glu	Phe	Phe	Lys	Asn	Val	Asn	Val	Ile	Glu	Asp	50	55	60	
Lys	Glu	Arg	Leu	Trp	Lys	Cys	Phe	Arg	Leu	Asn	Asp	Ser	Lys	Asn	Thr	65	70	75	80

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 506 to nucleotide 643;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 471 to nucleotide 765;
  - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AA35\_2 deposited under accession number ATCC 98303;
  - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;
  - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AA35\_2 deposited under accession number ATCC 98303;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
  - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
  - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
  - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.
3. A host cell transformed with a composition of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by a composition of claim 2, which process comprises:
  - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
  - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
  - (a) the amino acid sequence of SEQ ID NO:2;
  - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32;
  - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone AA35\_2 deposited under accession number ATCC 98303;the protein being substantially free from other mammalian proteins.
9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 32.
11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.
12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.
14. A composition comprising an isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 71 to nucleotide 736;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 113 to nucleotide 736;
  - (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 343;
  - (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303;
  - (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;
  - (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM42\_3 deposited under accession number ATCC 98303;
  - (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;
  - (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
  - (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;
  - (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
  - (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
  - (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).



15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- (b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 91;
- (c) fragments of the amino acid sequence of SEQ ID NO:4; and
- (d) the amino acid sequence encoded by the cDNA insert of clone AM42\_3 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 55 to nucleotide 423;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG137\_7 deposited under accession number ATCC 98303;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG137\_7 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:6;
- (b) the amino acid sequence of SEQ ID NO:6 from amino acid 62 to amino acid 123;
- (c) fragments of the amino acid sequence of SEQ ID NO:6; and
- (d) the amino acid sequence encoded by the cDNA insert of clone BG137\_7 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 186 to nucleotide 2030;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 873 to nucleotide 2030;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 802 to nucleotide 1173;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CH699\_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 218 to amino acid 329;

(c) fragments of the amino acid sequence of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone CH699\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 111 to nucleotide 677;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 156 to nucleotide 677;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO851\_1 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 120 to amino acid 189;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO851\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:10, SEQ ID NO:9 or SEQ ID NO:12 .
26. A composition comprising an isolated polynucleotide selected from the group consisting of:
- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 123 to nucleotide 755;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 279 to nucleotide 755;
  - (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 631;
  - (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP111\_1 deposited under accession number ATCC 98303;
  - (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;
  - (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP111\_1 deposited under accession number ATCC 98303;
  - (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;
  - (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
  - (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity;
  - (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
  - (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
  - (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 171;
- (c) fragments of the amino acid sequence of SEQ ID NO:14; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CP111\_1 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 214 to nucleotide 2760;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 406 to nucleotide 2760;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2011 to nucleotide 2565;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CS278\_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:16;
  - (b) the amino acid sequence of SEQ ID NO:16 from amino acid 596 to amino acid 784;
  - (c) fragments of the amino acid sequence of SEQ ID NO:16; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone CS278\_1 deposited under accession number ATCC 98303;
- the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 901 to nucleotide 1074;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 970 to nucleotide 1074;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 626 to nucleotide 1147;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303;

- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DF968\_3 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) fragments of the amino acid sequence of SEQ ID NO:18; and
- (c) the amino acid sequence encoded by the cDNA insert of clone DF968\_3 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 560 to nucleotide 820;



(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DN1120\_2 deposited under accession number ATCC 98303;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DN1120\_2 deposited under accession number ATCC 98303;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 61;

(c) fragments of the amino acid sequence of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone DN1120\_2 deposited under accession number ATCC 98303;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 649 to nucleotide 786;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 736 to nucleotide 786;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:21 from nucleotide 525 to nucleotide 787;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO589\_1 deposited under accession number ATCC 98303;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO589\_1 deposited under accession number ATCC 98303;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:22;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:22 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

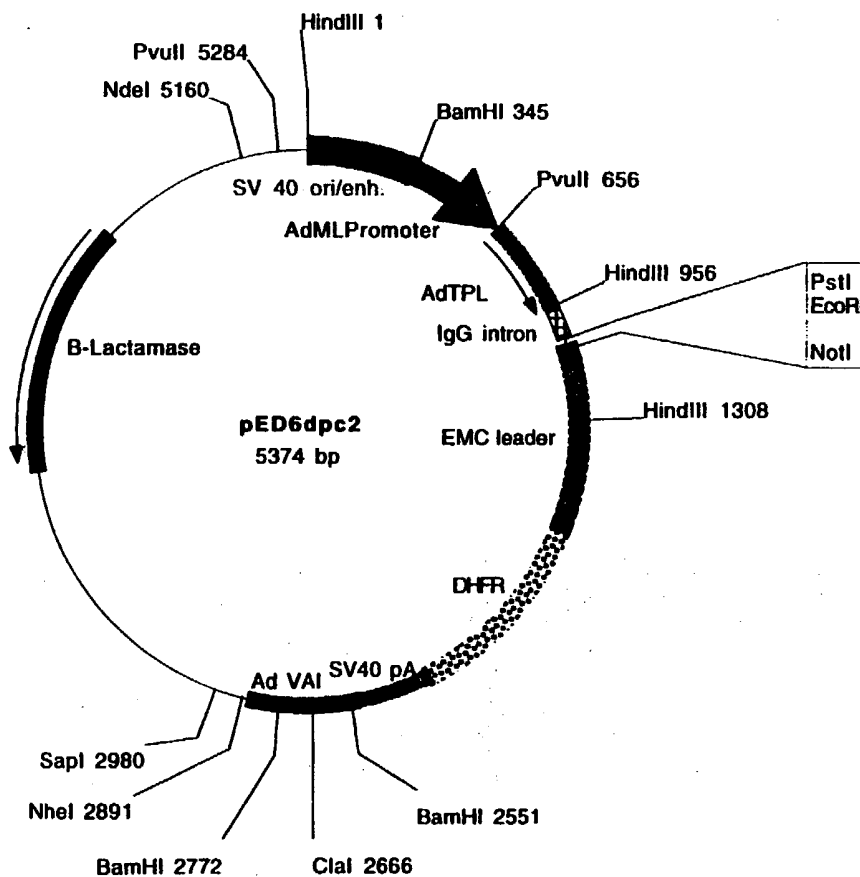
39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:22;

- (b) fragments of the amino acid sequence of SEQ ID NO:22; and
  - (c) the amino acid sequence encoded by the cDNA insert of clone DO589\_1 deposited under accession number ATCC 98303;
- the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:21.

FIGURE 1A

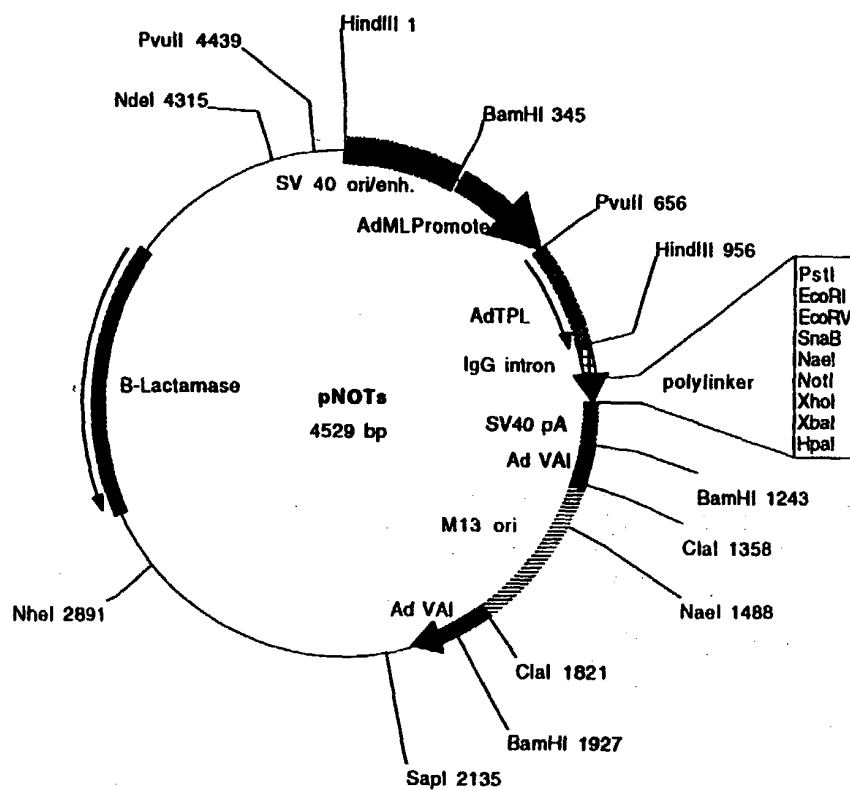


**Plasmid name:** pED6dpc2

**Plasmid size:** 5374 bp

**Comments/References:** pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



**Plasmid name:** pNOTs

**Plasmid size:** 4529 bp

**Comments/References:** pNOTs is a derivative of pMT2 (Kaufman et al, 1989. *Mol. Cell. Biol.* 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the Clal site. SST cDNAs are cloned between EcoRI and NotI